ELSEVIER

Journal of Chromatography B, 729 (1999) 65–74

JOURNAL OF CHROMATOGRAPHY B

Capillary zone electrophoresis and MALDI–mass spectrometry for the monitoring of in vitro *O*-glycosylation of a threonine/serine-rich MUC5AC hexadecapeptide

Benoît Soudan[a,*], Sylviane Hennebicq[a], Daniel Tetaert[a], Arnold Boersma[a], Colette Richet[a], Dominique Demeyer[a], Gilbert Briand[b], Pierre Degand[a]

[a]*Unité INSERM No. 377, Bâtiment G. Biserte, Place de Verdun, 59045 Lille Cédex, France*
[b]*Service Commun de Spectrométrie de masse, Université de Lille II, Faculté de Médecine, Place de Verdun, 59045 Lille Cédex, France*

Received 26 May 1998; received in revised form 9 February 1999; accepted 17 March 1999

Abstract

The in vitro *N*-acetylgalactosaminylation by human gastric UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases was assessed using the peptide motif GTTPSPVPTTSTTSAP, which is found naturally in the tandem repeat domains of the apomucin encoded by the gene *MUC5AC*. This peptide appeared to be an excellent tool for obtaining an insight into the extensive *O*-glycosylation processes of apomucins. Up to six *N*-acetylgalactosamines were added and the given glycopeptide species were well separated by capillary zone electrophoresis. Moreover, the degree of glycosylation (number of monosaccharide *O*-linked attachments) could be determined by MALDI–mass spectrometry without prior separation. Using different incubation times, we evidenced the accumulation of various glycopeptides, suggesting that the total glycosylation of an apomucin-peptide requires orderly *N*-acetylgalactosaminylation processing. This information was completed by experimental data showing that *N*-acetylgalactosaminylated octapeptides (the peptide backbones of which are part of GTTPSPVPTTSTTSAP) were able to selectively inhibit some *N*-acetylgalactosaminyltransferases. Our results suggest that this inhibition may influence the quality of the intermediate products appearing during the in vitro *O*-glycosylation process.

Keywords: *O*-Glycosylation; UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase

1. Introduction

The mucins, a major component of the mucus, are high-molecular-mass (up to several thousand kDa) glycoproteins characterised by a strong heterogeneity involving both the apomucin and the oligosaccharide chains [1,2]. To date, nine genes of epithelial mucins (named *MUC* in human) have been identified [3]. Deduced from the analysis of cDNA nucleotide sequences, the mucins share structural features with the model proposed for the mucin encoded by the gene *MUC2* [4]: i.e. (i) a peptide backbone with extended arrays of tandemly repeated (TR) sequences having a high percentage of threonine and/or serine, alanine and proline, (ii) poorly glycosylated cysteine-rich domains separating the tandem repeat domains and (iii) oligosaccharide chains in

*Corresponding author. Fax.: +33-03-2053-8562.
E-mail address: soudan@lille.inserm.fr (B. Soudan)

0378-4347/99/$ – see front matter
PII: S0378-4347(99)00124-3

O-glycosidic linkages through *N*-acetylgalactosamine (GalNAc) to the threonine or serine residues.

In view of a better understanding of the first step of mucin carbohydrate chain elongation, our group studies the *N*-acetylgalactosaminylation of peptide acceptors representative of TR domains deduced from *MUC* genes. *MUC5AC* is a gene that is expressed mainly in the respiratory and gastric tracts [5]. Up to now, analysis of the apomucin sequence encoded by *MUC5AC* has revealed that two octapeptides are frequently found in the TR domain, i.e., TTSTTSAP and GTTPSPVP (amino acids expressed using the one-letter code) [6]. The GTTPSPVP peptide is a better substrate for *N*-acetylgalactosaminylation in vitro by human gastric UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases than the most representative TR peptide, TTSTTSAP [7]. We describe in this work the in vitro glycosylation by human gastric microsomal preparations of a hexadecapeptide with the sequence GTTPSPVPTTSTTSAP, which is naturally found in the *MUC5AC* apomucin and corresponding to the juxtaposition of the two most frequent TR octapeptides in this apomucin. Capillary zone electrophoresis [8] and matrix-assisted laser desorption mass spectrometry (MALDI–MS) [9] appeared to be very powerful technologies, not only to discriminate the different glycopeptides obtained (up to six) but also to follow their formation during the incubation. Moreover, using three octapeptides that are constitutive of the peptide GTTPSPVPTTSTTSAP, i.e., GTTPSPVP, PSPVPTTS and TTSTTSAP, we evidenced that they were able to influence the *O*-glycosylation of the hexadecapeptide, in their glycosylated states.

2. Experimental

2.1. Materials

The Dowex resin AG 1-X8 (100–200 mesh, Cl^- form) was purchased from BioRad (Vitry s/Seine, France). 2,5-Dihydroxybenzoic acid (DHB) was from Aldrich (St. Quentin Fallavier, France). Uridine diphosphate *N*-acetyl-D-galactosamine, [galactosamine-1-^{3}H(N)]- (UDP-GalNAc) (8.7 Ci/mmol) and uridine diphosphate *N*-acetyl-D-galactosamine, [galactosamine-1-^{14}C] (54.7 mCi/mmol) were from New England Nuclear Chemicals (DuPont de Nemours, Les Ulis, France). MES [2-(*N*-Morpholino)ethanesulfonate], $MnCl_2$, PMSF (phenylmethanesulfonyl fluoride), AMP (adenosine 5′-monophosphate), E-64 [*trans*-epoxysuccinyl leucylamino (4-guanido) butane], ethylenediaminetetraacetic acid (EDTA), cold UDP-GalNAc, trifluoroacetic acid (TFA) and Triton X-100 were from Sigma (St. Louis, MO, USA). Polyvinyl alcohol (M_r 15 000) was from Fluka (Buchs, Switzerland). The synthetic peptide substrates NH_2-GTTPSPVPTTSTTSAP-COOH, NH_2-GTTPSPVP-COOH, NH_2-PSPVPTTS-COOH and NH_2-TTSTTSAP-COOH were obtained from Neosystem (Strasbourg, France). The purity of the peptides (>95%) was assessed by high-performance liquid chromatography analysis and capillary electrophoresis, and their masses were verified by electrospray mass spectrometry [10]. The amino acid sequences were also checked using an Applied Biosystems gas-phase Sequencer 477A.

2.2. Analytical methods

2.2.1. Measurement of radioactivity

The [^{3}H]- and [^{14}C]-labelled *N*-acetyl-D-galactosamine incorporated in the peptide substrates was measured by scintillation counting on a LS 3801 liquid scintillation system (Beckman Instruments, Irvine, CA, USA).

2.2.2. Capillary zone electrophoresis

Capillary zone electrophoresis (CZE) was performed as previously described [8] on a P/ACE system, model 5000 (Beckman, Fullerton, CA, USA) controlled by the system Gold software V810 (Beckman). The samples were introduced by nitrogen pressure at the positive electrode at a concentration of 2 mg/ml in 2 *M* formic acid containing 2.5% (v/v) polyvinyl alcohol [M_r 15 000] and were detected by UV spectrometry at 200 nm. The voltage across the capillary (57 cm length×75 μm internal diameter) was maintained at 15 kV and the cartridges were cooled to 23°C for all runs. The column was regenerated between each run by applying 0.1 *M*

NaOH for 2 min and pure water for 2 min in order to obtain reproducible migrations.

2.2.3. Mass spectrometry

MALDI–MS was adapted from the conditions described by Jespersen et al. [11]. A Vision 2000 time-of-flight instrument (Finnigan MAT, Bremen, Germany) equipped with a nitrogen laser operating at a wavelength of 337 nm was used. The instrument was operated in the reflector mode, the ions were accelerated to a potential of 6 kV in the ion source and post-accelerated by a conversion dynode in front of the detector to a potential of 10 kV. All spectra were obtained in the positive-ion mode and single-shot spectra were accumulated in order to obtain a good signal-to-noise ratio. The spectra were first externally calibrated with angiotensin and bovine insulin. The matrix, 2,5-dihydroxybenzoic acid (DHB), was dissolved in 0.1% trifluoroacetic acid (TFA) to a concentration of 10 mg/ml. The incubation products were diluted with matrix to a concentration of 2–8 pmol/μl and 1 μl of this sample-matrix solution was deposited on the target and dried under a cold air-stream before loading into the mass spectrometer.

2.3. Analytical procedures

2.3.1. Enzyme preparation

A macroscopically normal fundic human gastric sample (about 14 g) was obtained from a patient after surgery. The microsome suspension was prepared by Potter-Elvehjem homogenisation of the tissue, as previously described [10]. The microsomal pellet was suspended in a 0.2 *M* NaCl–0.25 *M* sucrose solution to obtain a final protein concentration of 5.5 g/l, and aliquots were stored at −80°C until use. The protein concentration was determined according to the method of Peterson [12].

2.3.2. Qualitative analysis of GalNAc incorporation in the GTTPSPVPTTSTTSAP peptide

In order to gain qualitative insight into the GalNAc incorporation process, the GTTPSPVPTTSTTSAP peptide was glycosylated for 0.5, 1, 2, 3, 5, 8 and 24 h and CZE analysis was performed on the unseparated glycosylated products. The assay mixtures in a total volume of 40 μl contained the following components, at final concentrations of: 1 m*M* GTTPSPVPTTSTTSAP (10 μl); 125 m*M* MES buffer (pH 7.2) (10μl) containing 0.2% (v/v) Triton X-100, 12.5 m*M* $MnCl_2$, 1 m*M* E-64, 1 m*M* PMSF; 1.25 m*M* AMP (5 μl); 1 m*M* cold UDP-*N*-acetylgalactosamine diluted with 1.25 n*M* tritiated UDP-GalNAc ($0.15–0.2\times10^6$ dpm; 5μl) and 15 μg of microsomal preparation (10 μl). After the addition of the enzyme preparation, samples were incubated in duplicate at 37°C. The reaction was stopped by the addition of eight volumes (320 μl) of 20 m*M* sodium borate–1 m*M* EDTA (pH 9.1). The reaction products were passed through AG 1-X8 resin (2 ml), eluted by 3 ml water and the incorporation rates were calculated by scintillation counting of aliquots of the samples with subtraction of values obtained from controls incubated without peptide substrate. Desalting of samples before CZE and MALDI–MS was performed as previously described [13].

2.3.3. N-Acetylgalactosaminylation of GTTPSPVPTTSTTSAP in the presence of glycosylated or unglycosylated octapeptides

To gain some insight into the GTTPSPVPTTSTTSAP glycosylation sites that are preferentially recognised by the *N*-acetylgalactosamine transferases, we attempted to *O*-glycosylate the peptide GTTPSPVPTTSTTSAP in the presence of potential competitive inhibitors, such as constituent octapeptides or glycosylated peptides, on which the glycosylation sites were previously determined [7]. The effects of these co-incubation assays on the appearance of the products corresponding to one- to six GalNAc *O*-linked hexadecapeptides was determined by CZE and MALDI–MS.

The unglycosylated peptides were obtained by chemical synthesis and the glycosylated peptides were prepared in the laboratory.

2.3.3.1. Preparative N-acetylgalactosaminylation of GTTPSPVP, PSPVPTTS and TTSTTSAP. The octapeptides GTTPSPVP, PSPVPTTS and TTSTTSAP (at 3 m*M*) were *N*-acetylgalactosaminylated under the conditions described previously, with 1 m*M* [^{14}C]-UDP-GalNAc (340 000 dpm for a total volume of 800 μl) by the same gastric microsomal homoge-

nate for 24 h. Each preparation was chromatographed on two AG 1-X8 resin columns (6 ml) and eluted with 12 ml of water. After freeze-drying, the samples were desalted on two C_{18} Sep-Pack cartridges [13], and the resulting products were pooled.

2.3.3.2. Glycosylation of GTTPSPVPTTSTTSAP in the presence of glycosylated or unglycosylated octapeptides. The peptide GTTPSPVPTTSTTSAP (1 m*M*) was *N*-acetylgalactosaminylated in duplicate with [^{3}H]-UDP-GalNAc for 24 h using the gastric microsomal preparation in the presence of 3- or 8 m*M* octapeptides (GTTPSPVP, PSPVPTTS or TTSTTSAP) or the glycosylated products of GTTPSPVP (2.5 m*M*), PSPVPTTS (0.3 m*M*) or TTSTTSAP (3.3 m*M*). The glycopeptide concentrations were expressed in term of m*M* GalNAc (this parameter was determined by the measure the incorporation of [^{14}C]-UDP-GalNAc in the glycosylated preparations after the desalting step). After 24 h incubation at 37°C, the reaction products were chromatographed on AG 1-X8 resin.

3. Results

3.1. Qualitative analysis of GalNAc incorporation in the GTTPSPVPTTSTTSAP peptide

The CZE profile of the glycosylation products obtained after 24 h reaction is given in Fig. 1. Identification of the peaks corresponding to the residual peptide and to the glycosylated products was achieved by comparing the profiles of the reaction products obtained either in the presence or absence of UDP-GalNAc (data not shown). Consequently, the peak P was attributed to the parent peptide and the peaks G1–G6 to glycopeptides. This interpretation was supported by previous experiments which showed that the addition of a GalNAc residue to a peptide or glycopeptide induced a longer migration time on CZE [7,8,13]. Hence, the G1 product(s) may correspond to the one *O*-GalNAc-peptide(s), the G2 product(s) to the two *O*-GalNAc-peptide(s)...and the G6 product(s) to the six *O*-GalNAc-peptide(s).

Another approach for the identification of the CZE glycosylated products was ESI–MS and MALDI–MS on the unseparated glycosylated preparations. By the electrospray mass spectrometry (ESI–MS) method, only three *m*/*z* products that corresponded to the parent peptide and to the mono- and di-GalNAc GTTPSPVPTTSTTSAP, respectively, were detected (data not shown) compared to the seven CZE peaks. By contrast, the MALDI–MS (Fig. 2) displayed *m*/*z*

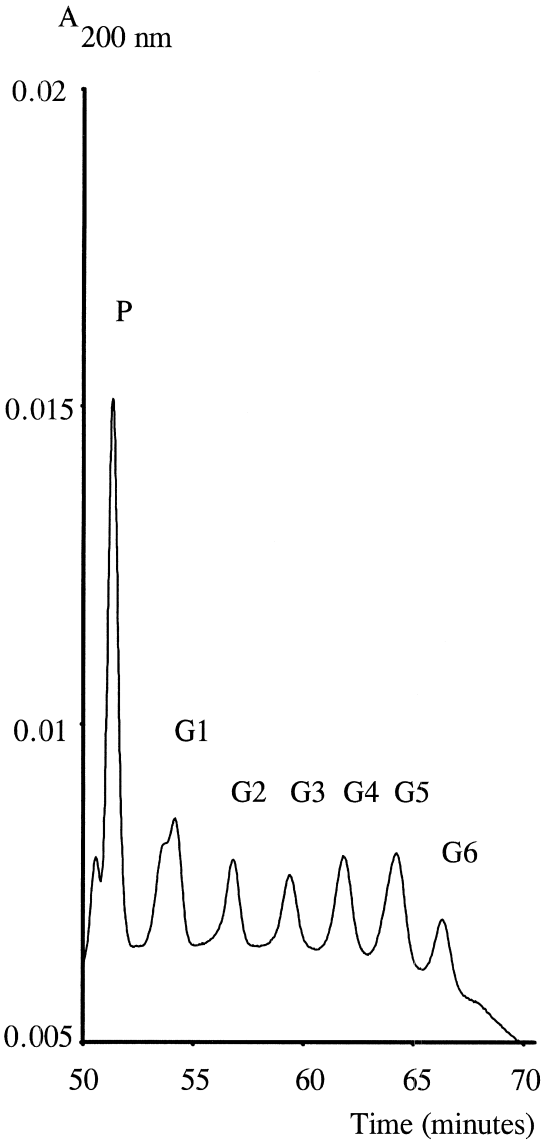

Fig. 1. Capillary zone electropherogram obtained from the 24-h glycosylation products of the GTTPSPVPTTSTTSAP peptide. P corresponds to the parent peptide and G1 to G6 to the glycopeptides.

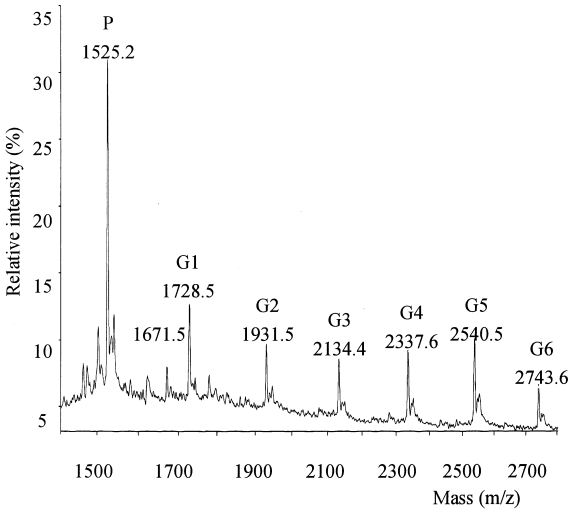

Fig. 2. Matrix-assisted laser desorption ionization mass spectrum of GTTPSPVPTTSTTSAP 24 h glycosylation products. The m/z values of 1525.2, 1728.5, 1931.5, 2134.4, 2337.6, 2540.5 and 2743.6 correspond to the parent peptide (P), and to the 1(G1), 2(G2), 3(G3), 4(G4), 5(G5) and 6 GalNAc (G6) substituted peptides associated with a sodium ion, respectively. m/z 1671.5 corresponds to the G1 glycopeptide(s) without the N-terminal glycine.

at 1525.2, 1728.5, 1931.5, 2134.4, 2337.6, 2540.5 and 2743.6 that corresponded to the parent peptide (P) $[M+Na^+]^+$ and to the 1 (G1), 2 (G2), 3 (G3), 4 (G4), 5 (G5) and 6 (G6) *O*-GalNAc-substituted peptides, respectively. Finally, the assignment of CZE fractions G1–G6 to the corresponding glycopeptides [e.g. G1 to the mono *O*-GalNAc glycopeptide(s)] was confirmed by the similarity between the CZE peak area relative percentages and the MALDI–MS signal intensity relative percentages: (respectively 19.5, 14.6, 14.2, 18.4, 22.2 and 11.1% for the glycopeptide CZE peaks, and 22.8, 15.3, 13.2, 17.5, 20.6 and 10.6% for the glycopeptide MALDI–MS signal intensities). Moreover, the CZE fraction with the label G1 appeared to be a double peak. The MALDI–MS of the unseparated reaction mixture revealed a m/z at 1671.5. Thus, it is likely that the shoulder before the CZE peak G1 corresponds to the G1 glycopeptide(s) without the N-terminal glycine. However, we cannot totally exclude the possibility that this peak may also be representative of different mono-GalNAc positional isomers with slightly different migration times.

3.2. N-Acetylgalactosaminylation kinetics of the peptide GTTPSPVPTTSTTSAP

Identification of the glycosylated peptide peaks on CZE was confirmed by the order of appearance of the CZE peaks during the kinetics of *N*-acetylgalactosaminylation of the peptide GTTPSPVPTTSTTSAP. In Fig. 3, the CZE G1 glycopeptide peak area was assigned a value of 100% for each incubation time and the areas of the

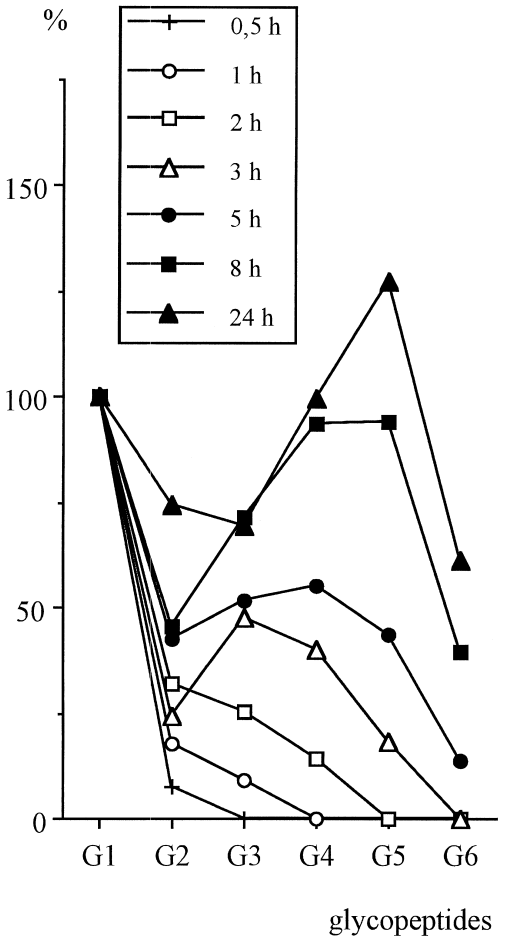

Fig. 3. Kinetics of formation of the glycopeptides during the *N*-acetylgalactosaminylation of GTTPSPVPTTSTTSAP. G1, G2, G3, G4, G5 and G6 correspond to the 1, 2, 3, 4, 5 and 6-GalNAc substituted peptides, respectively. The glycopeptide percentages were calculated by assigning a 100% value to the G1 CZE peak area for each incubation time and by determining the ratios of Gn/G1.

other CZE glycopeptide peaks were expressed as G*n*/G1 ratios. We observed the rapid formation of G1- and G2-glycopeptide(s) after a 30-min incubation period, while G3-glycopeptides appeared after 1 h. The G4-glycopeptides only appeared after a 2-h incubation period, followed by G5-glycopeptides after 3 h and, finally, G6-glycopeptides after 5 h. In addition, we evidenced a progressive and significant accumulation of G3- (after 3 h incubation), then G4- (after 5 h incubation) and G5-glycopeptides (after 8 h incubation).

3.3. N-Acetylgalactosaminylation of GTTPSPVPTTSTTSAP in the presence of glycosylated and unglycosylated octapeptides

3.3.1. Characterisation of the glycosylated octapeptide preparations

After AG 1-X8 chromatography, the total 24 h GalNAc incorporation was 500–600 nmol for GTTPSPVP and TTSTTSAP and 80 nmol for PSPVPTTS. The CZE profiles, confirmed by the MALDI–MS analyses of the octapeptide glycosylated preparations, revealed that the GTTPSPVP and TTSTTSAP octapeptides gave mono- and di-*N*-acetylgalactosaminylated peptides, whereas the PSPVPTTS octapeptide was only mono *N*-acetylgalactosaminylated. The sites of glycosylation were described previously [7]. Moreover, analysis of the glycosylated preparations revealed that they still contained a proportion of unglycosylated octapeptides. Indeed, the ratio between CZE peak areas of unglycosylated peptides and glycosylated products was 0.05 for the GTTPSPVP preparation, 0.95 for the TTSTTSAP preparation and 26.6 for the PSPVPTTS preparation. These ratios illustrated the different affinity(ies) of the GalNAc transferase(s) for the three octapeptides.

3.3.2. Glycosylation of GTTPSPVPTTSTTSAP in the presence of glycosylated or unglycosylated peptides

The level of incorporation of GalNAc was not significantly modified by co-incubation with the unglycosylated octapeptides at 3- or 8 m*M* (Fig. 4). In contrast, an important inhibition was found when the hexadecapeptide (at 1 m*M*) was co-incubated with the glycosylated octapeptide (GTTPSPVP, PSPVPTTS and TTSTTSAP) preparations at 2.5, 0.3 and 3.3 mM, respectively. The PSPVPTTS glycosylated preparation appeared to have the strongest inhibitory effect and this effect was indeed relevant to the glycopeptides, since the same co-incubation experiment in the presence of up to 8 m*M* unglycosylated peptide could not produce such an inhibition.

3.3.3. Quantitative and qualitative analysis of the GTTPSPVPTTSTTSAP glycosylated products by CZE

CZE was carried out on each co-incubation product. For the quantitative analysis, we calculated the relative percentages of the residual unglycosylated GTTPSPVPTTSTTSAP and glycopeptide species (G1 to G6) on the basis of their respective CZE peak areas (Table 1). For a qualitative analysis of the hexadecapeptide glycopeptide distribution, in order to favour the interpretation, the ratios of the G-peak areas were calculated proportionally to the mono-glycosylated peptide(s) (G1) (Table 2).

The unglycosylated octapeptides quantitatively inhibited the glycosylation of GTTPSPVPTTSTTSAP, but this inhibition was similar for each hexadecapeptide glycopeptide species (Gs) under each condition (Table 1). Moreover, Table 2 shows that the G*n*/G1 ratios remained unchanged when they were compared with the ratios obtained on incubation of GTTPSPVPTTSTTSAP alone.

The glycosylated octapeptides inhibited the appearance of the hexadecapeptide glycopeptide (Gs) CZE peaks (Table 1). However, we observed that this inhibition was not identical on each of the six potential glycopeptides, with the appearance of some glycopeptide species being totally inhibited. We evidenced that the GTTPSPVP glycopeptides had a lesser inhibitory effect on the first glycosylation of GTTPSPVPTTSTTSAP than the PSPVPTTS and TTSTTSAP ones. Moreover, the GTTPSPVP glycopeptides had a marked negative effect on the production of G4 and totally inhibited the formation of G5 and G6 (Table 1). The PSPVPTTS and TTSTTSAP glycopeptides not only had a strong inhibitory effect on the first glycosylation step (G0 to G1), but also completely inhibited the formation of G4, G5 and G6. Moreover, the PSPVPTTS glyco-

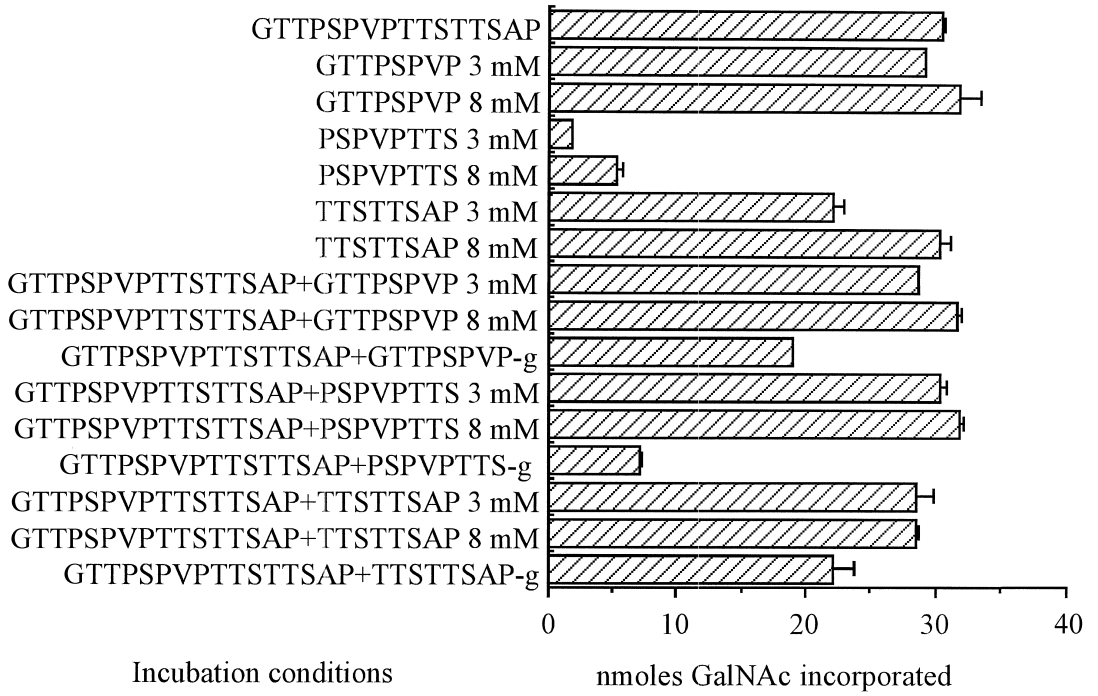

Fig. 4. Incorporation of GalNAc (±SEM) on the peptide substrates after a 24-h incubation period in the presence of 1 m*M* UDP-GalNAc in a total volume of 40 μl. The GTTPSPVPTTSTTSAP peptide was always used at a concentration of 1 m*M*. The unglycosylated peptides GTTPSPVP, PSPVPTTS and TSTTSAP were used at 3- or 8 m*M*, and the glycosylated products of these octapeptides (GTTPSPVP-g, PSPVPTTS-g and TTSTTSAP-g) were used at 2.5, 0.3 and 3.3 m*M*, respectively.

peptides markedly decreased the G1 to G2 transition and the TTSTTSAP glycopeptides decreased the G2 to G3 transition (Table 1).

Qualitative analysis of the GTTPSPVPTTSTTSAP glycopeptide distribution (Table 2) revealed that the GTTPSPVP glycopeptides induced a significant de-

Table 1
Quantitative distribution of the unglycosylated GTTPSPVPTTSTTSAP (P) and the glycosylated hexadecapeptides (G1 to G6) under different 24-h incubation conditions with a gastric microsomal preparation[a]

Conditions of incubation	P (%)	Glyco-peptides (%)	G1 (%)	G2 (%)	G3 (%)	G4 (%)	G5 (%)	G6 (%)
GTTPSPVPTTSTTSAP (control)	85.57	14.43	5.37	2.95	1.7	1.97	1.90	0.54
GTTPSPVPTTSTTSAP+GTTPSPVP, 3 m*M*	91.78	8.22	3.33	1.89	0.96	0.97	0.89	0.18
GTTPSPVPTTSTTSAP+PSPVPTTS, 3 m*M*	88.03	11.97	4.57	2.28	1.37	1.58	1.66	0.51
GTTPSPVPTTSTTSAP+TTSTTSAP, 3 m*M*	91.6	8.4	3.08	2.09	0.77	0.95	1.19	0.32
GTTPSPVPTTSTTSAP+GTTPSPVP-glycopeptides, 2.5 m*M*	93.44	6.56	3.85	1.66	0.77	0.28	0	0
GTTPSPVPTTSTTSAP+PSPVPTTS-glycopeptides, 0.3 m*M*	97.01	2.99	1.99	0.59	0.41	0	0	0
GTTPSPVPTTSTTSAP+TTSTTSAP-glycopeptides, 3.3 m*M*	96.66	3.34	2.06	1.05	0.33	0	0	0

[a] These percentages were determined by comparison of the CZE peak areas. The hexadecapeptide GTTPSPVPTTSTTSAP was always used at a concentration of 1 m*M*.

Table 2
Qualitative distribution of the GTTPSPVPTTSTTSAP glycopeptide CZE peak areas under different 24 h incubation conditions with a gastric microsomal preparation[a]

Conditions of incubation	Ratio G1/G1	Ratio G2/G1	Ratio G3/G1	Ratio G4/G1	Ratio G5/G1	Ratio G6/G1
GTTPSPVPTTSTTSAP	1	0.55	0.31	0.36	0.37	0.09
GTTPSPVPTTSTTSAP+GTTPSPVP 3mM	1	0.57	0.29	0.29	0.26	0.05
GTTPSPVPTTSTTSAP+PSPVPTTS 3 mM	1	0.51	0.29	0.34	0.35	0.11
GTTPSPVPTTSTTSAP+TTSTTSAP 3 mM	1	0.68	0.25	0.30	0.38	0.10
GTTPSPVPTTSTTSAP+GTTPSPVP-glycopeptides 2.5 mM	1	0.43	0.19	0.07	0	0
GTTPSPVPTTSTTSAP+PSPVPTTS-glycopeptides 0.3 mM	1	0.29	0.20	0	0	0
GTTPSPVPTTSTTSAP+TTSTTSAP-glycopeptides 3.3 mM	1	0.51	0.16	0	0	0

[a] The hexadecapeptide GTTPSPVPTTSTTSAP was always used at a concentration of 1 m*M*.

crease in the ratio G4/G1 and the PSPVPTTS and TTSTTSAP glycopeptides caused a decrease in the ratios of G2/G1 and G3/G1, respectively.

4. Discussion

The peptide GTTPSPVPTTSTTSAP, encoded by the gene *MUC5AC*, is of great interest because it is multiglycosylated in vitro. Thereby, this peptide may be a good tool to investigate the rules responsible for an orderly *N*-acetylgalactosaminylation of the apomucins. Nine *N*-acetylgalactosaminylation sites are available on our peptide substrate, i.e., six threonines and three serines. However, we observed only the *O*-linkage of six GalNAc on the peptide.

Electrospray mass spectrometry, which was of great aid in the study of octapeptide glycosylation [7], did not allow the detection in the unseparated reaction mixture of more than two GalNAc *O*-linkages on GTTPSPVPTTSTTSAP, while the CZE profiles revealed clearly the formation during the incubation of up to six glycopeptides corresponding to one to six GalNAc added to the peptide backbone (G1 to G6). In contrast, by MALDI–MS, the six glycopeptides were identified and their qualitative distribution was comparable to that observed using CZE. These results confirmed that MALDI–MS was a better procedure to analyze complex mixtures than ESI–MS and was not so sensitive to contaminants.

Unfortunately, the identification of the exact GalNAc *O*-linkage sites was unsuccessful whatever procedures used, i.e. (i) preparative CZE for Edman degradation of the purified glycopeptide fractions, because the CZE fractions were too close together to be used for further preparative purposes and gas-phase sequencing analysis; (ii) mass spectrometry techniques, such as post-source decay matrix-assisted laser desorption ionization (PSD–MALDI), which was performed on the unseparated reaction mixture. This latter technique was not reliable for determining the sites of sugar attachments on the peptide backbone since these sugars fell off the fragment ions and, therefore, their positions were not detectable. This seems to be due to the sequence of the *MUC5AC* peptide backbone, since good glycopeptide sequencing and localisation spectra were observed under the same experimental conditions for a *MUC*1 peptide. (S. Goletz, personal communication) [14,15]. Moreover, this mass spectrometry failure was also observed by all groups working on *O*-glycopeptide sequencing using different instrumentation [16].

The CZE procedure, using 2.5% polyvinyl alcohol (M_r 15 000) in the electrophoresis buffer solution [8], provided adequate conditions for a dynamic coating of the silica capillary walls that induced a good separation of the unglycosylated and glycosylated peptides. Moreover, this complete CZE separation of glycosylated products allowed us, during incubation times, to follow the evolution of the glycosylation of GTTPSPVPTTSTTSAP. A main result was that the ratios of the different glycopeptide CZE peaks (G1 to G6) were modified over time, with a progressive accumulation of G3 after 3 h incubation, G4 after 5 h and G5 after 8 h. Under these conditions, this accumulation may reflect the increase during the incubation time of some glyco-

sylated positional isomers that could not be further glycosylated. In fact, up to the 2-h time-point, the glycopeptide repartition indicated that the glycosylated products were naturally converted to the upper glycosylated homologues. Up to 3 h, G3 was partly further glycosylated but began to be accumulated, and this accumulation phenomenon was amplified afterwards, as it was for G4 and G5.

The positional isomers were probably relevant to the activity of different GalNAc transferase species: UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase(s) directly acting on apomucins, and UDP-GalNAc:glycopeptide *N*-acetylgalactosaminyltransferase(s) only acting on already *N*-acetylgalactosaminylated apomucins on determined positions [17]. The GalNAc transferase heterogeneity was confirmed by the results of the incubation of GTTPSPVPTTSTTSAP with glycosylated octapeptides, which were demonstrated to quantitatively and qualitatively inhibit the hexadecapeptide glycosylation, and this inhibition differed depending on the glycopeptide products used. Indeed, the GTTPSPVP glycopeptides probably inhibited GalNAc transferase(s) responsible for the appearance of G1, G3 and G4. The PSPVPTTS glycopeptides strongly inhibited GalNAc transferase(s) responsible for G1 and G2 positional isomers and the G3 produced isomers could not be more glycosylated. Finally, the TTSTTSAP glycopeptides also inhibited the GalNAc transferases responsible for the first step of the GTTPSPVPTSTTSAP glycosylation process (formation of G1). These G1 positional isomers were appropriate substrates for a diglycosylation of the hexadecapeptide, but the GalNAc transferase activity(ies) responsible for the G2 to G3 transition was/were decreased and there was no G4; this latter effect may also be relevant to the inhibition of other GalNAc transferase activity(ies). Likewise, the unglycosylated octapeptides quantitatively inhibited the first glycosylation of GTTPSPVPTTSTTSAP, but the glycopeptide repartition appeared to be unaffected afterwards.

In conclusion, the identification of positional isomers may be of great interest in understanding the order in which the apomucins are extensively *O*-glycosylated in vivo, but also to get insights into the *N*-acetylgalactosamine transferases that are responsible for the incomplete glycosylation of the same apomucins in epithelial cancer cells. Indeed, in tumoral processes, the mucin glycoprotein carbohydrate chains display modifications of peripheral structures but also the exposure of inner core regions, i.e., the Thomsen-Friedenreich antigens Tn (GalNAc-α-Thr/Ser) and T (Galβ1,3 GalNAc-α-Thr/Ser) [18]. Up to now, it is supposed that this elongation default of native oligosaccharide chains may be relevant to a random *N*-acetylgalactosaminylation of the apomucin backbone by the GalNAc transferases of tumoral cells, inducing inadequate substrates for glycosyltransferases. This hypothesis is supported by preliminary data showing that the UDP-galactose:glycoprotein-*N*-acetyl-D-galactosamine 3-β-D-galactosyltransferase (implicated in a second step in the glycosylation pathway of *O*-glycans) is controlled by structure and by initial glycosylation of the peptide core substrate [19]. Thereby, the elongation default of mucin glycoprotein carbohydrate chains in epithelial cancer cells may be due to the overexpression of some GalNAc transferases that preferentially recognise already *N*-acetylgalactosaminylated substrates at different glycosylation stages. Recent findings reported such UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase isoforms [17,20]. In tumoral processes, these glycopeptide transferases may be competitors with other *N*-acetylgalactosaminyltransferases that are normally implicated in the initiation of apomucin glycosylation.

Acknowledgements

We are indebted to Dr. G. Ricart (Laboratoire de Spectrométrie de Masse, Bat. C4, USTL, 59655 Villeneuve d'Ascq cedex, France) for technical support and help in the interpretation of MALDI–mass spectra and to Dr S. Goletz (Max-Delbrück-centrum für Molekulare Medizin, Robert-Rössle Strasse 10, D-13125 Berlin-Buch, Germany) for performing post-source decay matrix-assisted laser desorption ionization (PSD–MALDI).

References

[1] P. Roussel, G. Lamblin, in: J. Montreuil, J.F.G. Vliegenthart, H. Schachter (Eds.), Glycoproteins and Disease, Elsevier, Amsterdam, 1996, pp. 351–393.

[2] N.W. Toribara, A.M. Roberton, S.B. Ho, W.L. Kuo, E. Gum, J.W. Hicks, J.R. Gum, J.C. Byrd, B. Siddiki, Y.S. Kim, J. Biol. Chem. 268 (1993) 5879–5885.
[3] L.E. Vinall, A.S. Hill, P. Pigny, W.S. Pratt, N. Toribara, J.R. Gum, Y.S. Kim, N. Porchet, J.P. Aubert, D.M. Swallow, Hum. Genet. 102 (1998) 357–366.
[4] N.W. Toribara, J.R. Gum, P.J. Culhane, R.E. Lagace, J.W. Hicks, G.M. Petersen, Y.S. Kim, J. Clin. Invest. 88 (1991) 1005–1013.
[5] J.P. Audie, A. Janin, N. Porchet, M.C. Copin, B. Gosselin, J.P. Aubert, J. Histochem. Cytochem. 41 (1993) 1479–1485.
[6] V. GuyonnetDuperat, J.P. Audie, V. Debailleul, A. Laine, M.P. Buisine, S. Galiegue-Zouitina, P. Pigny, P. Degand, J.P. Aubert, N. Porchet, Biochem. J. 305 (1995) 211–219.
[7] S. Hennebicq, D. Tetaert, B. Soudan, A. Boersma, G. Briand, C. Richet, J. Gagnon, P. Degand, Glycoconjugate J. 15 (1998) 275–282.
[8] Y. Boulis, C. Richet, K. Haupt, S. Hennebicq, B. Soudan, D. Tetaert, P. Degand, M.A. Vijayalakshmi, J. Chromatogr. A 805 (1998) 285–293.
[9] M. Karas, U. Bahr, Mass Spectrom. Rev. 10 (1991) 335–357.
[10] D. Tetaert, B. Soudan, J.M. Lo-Guidice, C. Richet, P. Degand, G. Boussard, C. Mariller, G. Spik, J. Chromatogr. B 658 (1994) 31–38.
[11] S. Jespersen, J.A. Koedam, C.M. Hoogerbrugge, U.R. Tjaden, J. VanderGreef, J.L. VandenBrande, J. Mass Spectrom. 31 (1996) 893–900.
[12] G.L. Peterson, Anal. Biochem. 83 (1977) 346–356.
[13] S. Hennebicq-Reig, B. Soudan, D. Tetaert, G. Briand, C. Richet, D. Demeyer, P. Degand, Carbohydr. Res. 280 (1996) 121–126.
[14] S. Müller, S. Goletz, N. Packer, A. Gooley, A.M. Lawson, F.-G. Hanish, J. Biol. Chem. 272 (1997) 24780–24793.
[15] S. Goletz, B. Thiede, F.-G. Hanish, M. Schultz, J. Peter-Katalinic, S. Müller, O. Seitz, U. Karsten, Glycobiology 7 (1997) 881–896.
[16] K. Alving, R. Körner, H. Paulsen, J. Peter-Katalinic, J. Mass Spectrom. 33 (1998) 1124–1133.
[17] K.G. TenHagen, D. Tetaert, F.K. Hagen, C. Richet, T.M. Beres, J. Gagnon, M.M. Balys, B. VanWuyckhuyse, P. Degand, L.A. Tabak, Glycobiology 8 (1998) 1095.
[18] Y.S. Kim Jr., J. Gum, I. Brockhausen, Glycoconjugate J. 13 (1996) 693–707.
[19] M. Granovsky, T. Biefeldt, H. Paulsen, M. Meldal, J. Brockhausen, I. Brockhausen, Eur. J. Biochem. 221 (1994) 1039–1046.
[20] E.P. Bennett, H. Hassan, U. Mandel, E. Mirgorodskaya, P. Roepstorff, J. Burchell, J. Taylor-Papadimitriou, M.A. Hollingsworth, G. Merkx, A.G. VanKessel, H. Eiberg, R. Steffensen, H. Clausen, J. Biol. Chem. 273 (1998) 30472–30481.